RobinX: an XML-driven classification for round-robin sports timetabling

David Van Bulck · Dries Goossens · Jörn Schönberger · Mario Guajardo

Keywords Timetabling · Sports scheduling · Benchmark · XML · Validation

1 Introduction

Creating timetables for sports competitions is a complex matter due to conflicting interests of many stakeholders. Apart from some basic constraints, each competition has its own requirements and as a result, no generally applicable solution method exists [8]. For example, the minimization of consecutive home and away matches is key within most professional competitions (e.g. [1,4]), whereas respecting player availability is far more important in non-professional competitions (e.g. [6,14]).

In a round-robin timetable, often called a schedule, every team plays against every other team a fixed number of times. Although a few contributions have been made to organize various constraints that occur in round-robin timetabling, they did not result in a generally applicable classification or file format for problem instances and solutions. Bartsch, Drexl, and Kröger [1] describe several sports competitions by means of organizational, attractiveness, and fairness constraints. Rasmussen and Trick [13] list eight common constraint types. Kendall, Knust, Ribeiro, and Urrutia [8] are the only authors to distinguish between different objective functions. Nurmi et al. [10] are the first to set up a (plain text-only) file format to store instances, and propose a set of artificial and real-world instances together with the best solutions found so far.

D. Van Bulck
Ghent University, Faculty of Economics and Business Administration
E-mail: david.vanbulck@ugent.be

D. Goossens
Ghent University, Faculty of Economics and Business Administration

J. Schönberger
Technische Universität Dresden, Institute of Transport and Economics

M. Guajardo
NNH Norwegian School of Economics, Department of Business and Management Science

This format, however, has limited utility with respect to the ease of data manipulation and is not extensible towards several real-world problems. The lack of a generally accepted classification and data format makes it extremely difficult to assess algorithmic performance. This forms one of the main obstacles in current algorithmic progress since little general insights have been gained from previous computational studies. One notable exception is the traveling tournament problem [3] that minimizes the total team travel in the timetable. For this problem, various algorithmic results have been reported after a set of artificial benchmarking instances had been made publicly available [3].

This paper presents preliminary results of an XML-driven three-field classification scheme for round-robin sports-timetabling problems. For the definition of an instance and its solution, we propose two `XML`-file-based templates and a `C++`-library (RobinX) to evaluate both files[1].

2 Three-field notation

In the late 1970s, Graham et al. [7] introduced a three-field notation that is now widely used to distinguish between different machine scheduling problems. This section proposes a similar three-field notation ($\alpha/\beta/\gamma$) to describe different variants of the sports-timetabling problem.

The α-field represents the competition format (α_1), the compactness (α_2), and the symmetry properties of the timetable (α_3). First, the α_1 part denotes a k round-robin tournament between n teams by the symbol $\mathtt{kRR}^n$. Next, the α_2 part categorizes a timetable as compact ($\mathtt{C}^m$) if the number of available time slots m is no more than the minimal number required to play all matches [8]. Otherwise the timetable is time-relaxed ($\mathtt{R}^m$). Many k round-robin timetables are split into k intervals that each contain a single round robin. In the mirrored scheme (`M`), for example, the order of the opponents in each interval is identical to that of the previous interval but the home advantage is switched. Similarly, the α_3 field can take on one of the values `I` (inverse scheme), `E` (English scheme), or `F` (French scheme) [5].

Second, the β-field lists around 30 constraints partitioned into six constraint groups that classify the vast majority of the constraints from the literature. Place constraints (`PL`) enforce a team to play home or away in a certain time slot. Group constraints (`GR`) define restrictions applicable for a set of teams (team groups) or a set of time slots (slot groups). Next, break constraints (`BR`) limit and distribute the number of breaks, i.e. two successive home or away matches. Game constraints (`GA`) enforce or forbid a match to be assigned to a particular time slot. Fairness constraints (`FA`) distribute timetable inconveniences over all teams. Finally, separation constraints (`SE`) regulate the number of time slots between successive matches of a team. Constraints are either hard or soft and have a non-negative penalty weight: the total cost induced by a constrained is (a transformation of) the amount of violation multiplied with the penalty weight [9, 11]. The cost of violated hard constraints contributes to the infeasibility value of a timetable whereas the cost of violated soft constraints is included

[1] All source code and a more fine-grained description of this project can be found on `www.sportscheduling.ugent.be/RobinX`.

Table 1 Example of a specific constraint for each constraint group.

PL1	No team from team group T can have less than k_{min} or more than k_{max} home matches *(away matches, matches)* in slot group S.
GR3	At least n_{min} and at most n_{max} teams from team group T must simultaneously play a home match *(away match, match)* in slot group S.
BR1	No team from team group T can have a break in slot group S.
GA1	All matches *(the k-th match)* between *(home)* team i and team j must be timetabled before *(after)* time slot s.
FA2	The difference in the number of home matches *(away matches, matches)* for any pair of teams must not be larger than k in any *(the final)* stage of the tournament.
SE3	Any team in team group T must have at least m_{min} and at most m_{max} time slots between consecutive matches.

into the objective value of a timetable. A timetable is feasible if and only if it has an infeasibility value of zero.

Lastly, the γ-field refers to the objective function in use. The minimum break objective (BM) constructs break-minimal timetables. Alternatively, the minimum travel-distance function (TR) can be used to model the traveling tournament problem [3]. Based on estimated costs or revenues, the minimum cost or maximum revenue problem (CR) constructs cost-minimal timetables [2]. In contrast to the previous objectives, which require that all constraints from the α-field are hard, the soft-constraints objective (SC) minimizes the cost resulting from violated soft constraints while still respecting all hard constraints.

3 XML Format

Motivated by the success of XML in other research disciplines [9, 11, 12], we propose two intuitive XML standards for exchanging datasets in the field of sports timetabling. The first standard is used to store an instance of the problem and is made up of six different blocks. To begin with, a metadata block stores descriptive information such as the name of the contributor. Next, a data block holds static information, e.g. pairwise team distances, needed for the evaluation of a solution. Then, a resource block defines all team groups, teams, slot groups, and slots. The last three blocks encode the three-field notation as outlined in Section 2. The second XML standard represents a solution for the problem instance. This XML file consists of two blocks: a metadata block, and a block that enumerates all matches. In addition, we present RobinX: a free and open-source C++-library to read, write, generate, manipulate, validate, and evaluate both XML-files. If a solution respects all hard constraints as described by the instance file, the program returns the objective value of the solution. Otherwise, the program returns a list with all violated soft and hard constraints. RobinX is embedded in a user-friendly web application with a database of benchmark instances and their best-known solutions. This database uses the three-field notation to group problem instances and their solutions into several archives that can be searched through by a simple query tool. With this paper, we invite researchers to join the project and submit their own problem instances and solutions.

References

1. Bartsch, T., Drexl, A., Kröger, S.: Scheduling the professional soccer leagues of Austria and Germany. Comput. Oper. Res. **33**, 1907–1937 (2006)
2. Briskorn, D., Drexl, A.: A branching scheme for finding cost-minimal round robin tournaments. EJOR **197**, 68–76 (2009)
3. Easton, K., Nemhauser, G., Trick, M.: The traveling tournament problem description and benchmarks. In: T. Walsh (ed.) Principles and Practice of Constraint Programming — CP 2001, pp. 580–584. Springer, Berlin, Heidelberg (2001)
4. Goossens, D., Spieksma, F.: Scheduling the Belgian soccer league. Interfaces **39**, 109–118 (2009)
5. Goossens, D., Spieksma, F.: Soccer schedules in Europe: an overview. J. Sched. **15**, 641–651 (2011)
6. Goossens, D., Spieksma, F.: Indoor football scheduling. In: E. Özcan, E.K. Burke, B. McCollum (eds.) Proceedings of the 10th International Conference on the Practice and Theory of Automated Timetabling, pp. 167–178. PATAT (2014)
7. Graham, R.L., Lawler, E.L., Lenstra, J.K., Rinnooy Kan, A.H.G.: Optimization and approximation in deterministic sequencing and scheduling: a survey. Annals of discrete mathematics **5**, 287–326 (1979)
8. Kendall, G., Knust, S., Ribeiro, C.C., Urrutia, S.: Scheduling in sports: An annotated bibliography. Comput. Oper. Res. **37**, 1–19 (2010)
9. Kingston, J.H., Post, G., Vanden Berghe, G.: A unified nurse rostering model based on XHSTT. In: Proceedings of the 12th International Conference on the Practice and Theory of Automated Timetabling. PATAT, Vienna (2018)
10. Nurmi, K., Goossens, D., Bartsch, T., Bonomo, F., Briskorn, D., Durań, G., Kyngäs, J., Marenco, J., Ribeiro, C.C., Spieksma, F., Urrutia, S., Wolf, R.: A framework for scheduling professional sports leagues. In: H. Katagir, L. Xu, A.H.S. Chan (eds.) Ao,S-I., vol. 5, p. 1428. American Institute of Physics (2010)
11. Post, G., Ahmadi, S., Daskalaki, S., Kingston, J.H., Kyngäs, J., Nurmi, C., Ranson, D.: An XML format for benchmarks in high school timetabling. Ann. Oper. Res. **194**, 385–397 (2012)
12. Post, G., Kingston, J.H., Ahmadi, S., Daskalaki, S., Gogos, C., Kyngäs, J., Nurmi, C., Musliu, N., Pillay, N., Santos, H., et al.: XHSTT: an XML archive for high school timetabling problems in different countries. Ann. Oper. Res. **218**, 295–301 (2014)
13. Rasmussen, R.V., Trick, M.A.: Round robin scheduling–a survey. EJOR **188**, 617–636 (2008)
14. Schönberger, J., Mattfeld, D.C., Kopfer, H.: Memetic algorithm timetabling for non-commercial sport leagues. EJOR **153**, 102–116 (2004)